Bayesian Spatial Survival Models for Political Event Processes*

David Darmofal
Department of Political Science
University of South Carolina
350 Gambrell Hall
Columbia, SC 29208
Phone: (803) 777-5440
Email: darmofal@gwm.sc.edu

Abstract

Research in political science is increasingly, but independently, modeling heterogeneity and spatial dependence in political processes. This paper draws together these two research agendas via spatial random effects survival models. In contrast to existing survival models in political science, which assume spatial independence, spatial survival models allow for spatial autocorrelation in random effects at neighboring locations, which will occur if we are unable to model fully the sources of spatial autocorrelation in our data. I examine spatially dependent random effects in both semiparametric Cox and parametric Weibull models, and examine these random effects in both individual and hierarchical frailty models. I employ a Bayesian approach in which spatial autocorrelation in unmeasured risk factors across neighboring units is incorporated into the model via a conditionally autoregressive (CAR) prior. I apply the Bayesian spatial survival modeling approach to the timing of U.S. House members' position announcements on the North American Free Trade Agreement (NAFTA). I find that spatial shared frailty models outperform standard non-frailty models and non-spatial frailty models in both the semiparametric and parametric analyses. The modeling of the spatial dependence in the random effects also produces changes in the effects of substantive covariates.

*A previous version of this paper was presented at the 2006 Annual Meeting of the American Political Science Association. I thank Sudipto Banerjee, Katherine Barbieri, Neal Beck, Janet Box-Steffensmeier, Brad Carlin, Rob Franzese, Darren Greenwood, Jude Hays, Andrew Lawson, Lyndsey Young Stanfill, Tracy Sulkin, and Chris Zorn for helpful conversations on this project.

1 Introduction

Political events, by their nature, involve shared concerns and interdependent actors. As a consequence, the occurrence of a political event in one location is often associated with similar events in neighboring locations. Concepts such as the domino effect, waves of democratization, and policy diffusion highlight the spatial dimension in many political event processes (Huntington 1991; Berry and Berry 1992). But while concepts of spatial interaction and diffusion are central to many of our theories of event processes in political science, our approach to modeling this spatial dimension has often been quite limited. Generally, spatial dependence in survival data has been modeled via simple indicators such as the number or proportion of neighboring units that previously experienced the event of interest (Berry and Berry 1990; but see Berry and Baybeck 2005; see also Starr 1991). While useful in highlighting the interdependencies in political event processes, such an approach does not provide a generalized method for modeling spatial dependence nor is it consistent with the simultaneous nature of spatial, as opposed to temporal, dependence.

This standard approach to modeling spatial dependence in survival data contrasts with an emerging interest in modeling spatial dependence in other political science data (see, e.g., Cho 2003; Gimpel and Cho 2004; Beck, Gleditsch, and Beardsley 2006; Darmofal 2006; Franzese and Hays 2007). This latter literature applies spatial econometrics to model spatial dependence in continuous dependent variables. These studies demonstrate the importance of modeling spatial autocorrelation in order to avoid biased and inconsistent parameter estimates and biased standard errors and thus draw valid inferences about political phenomena.

As is well-known, standard econometric approaches developed for continuous dependent variables are not applicable for survival data. Right-censoring of time-to-event data produces observational equivalence between units experiencing the event at the end of the observation period and those censored and yet to experience the event. As a consequence, survival or event history models incorporating a censoring indicator are employed to model event processes (Box-Steffensmeier and Jones 2004).[1] The same complications that right-censoring produces in standard survival data also preclude the application of spatial econometric models developed for continuous dependent variables to spatially dependent survival data. As a consequence, scholars concerned about modeling the spatial dependence predicted by our theories of event processes in political science will wish to employ spatial survival models that account for spatial autocorrelation in units' risk propensity.

Recently, the non-spatial survival modeling literature has seen an increased interest in the use of random effects models to account for unobserved heterogeneity in units' risk propensity (Hill, Axinn, and Thornton 1993; Klein and Moeschberger 1997; Hougaard 2000). Here, scholars account for unmeasured sources of variation in risk propensity via either unit-specific (individual) or hierarchical (shared) random effects, or frailty terms. By incorporating such random effects, scholars can account for the fact that some units are more frail than others – and thus have a higher propensity to experience the event of interest – and avoid biased parameter estimates and violations of the proportional hazards assumption that are induced by a false assumption of homogeneity (see, e.g., Box-Steffensmeier and Jones 2004, 147-48). Testing for unobserved heterogeneity then simply involves estimating the variance term, θ, of the random effects. A significant positive value of θ indicates unmodeled heterogeneity in risk propensity while a value of θ that is indistinguishable from the null indicates that sources of variation in risk propensity are accounted for via the covariates in the model.

Recent years have seen considerable progress in the use of frailty models within political

[1] The term survival model derives from the study of patient mortality in biostatistics, an important field in the development of these models. Political scientists also refer to these models as survival models. Within political science, however, the interest is not in physical mortality, but instead in the survival of units until they experience a political event of interest. Thus, for example, political scientists speak of government or cabinet survival (Warwick 1992), conflict survival (Regan 2002), and survival in office (Jones 1994). Survival models are also known as event history and duration models.

science (Carpenter 2002; Gordon 2002; Chiozza and Goemans 2004; Colaresi 2004; Box-Steffensmeier and De Boef 2006). These studies demonstrate the importance of accounting for unobserved or unmeasured sources of heterogeneity in event occurrence. But while some of these studies examine dependence across competing risks (Gordon 2002) or repeated events (Box-Steffensmeier and De Boef 2006), all applications of frailty models in political science assume that the random effects are spatially independent. Given the theoretical predictions of spatial dependence in political event processes, political scientists will, I argue, often wish to allow for spatial dependence across random effects in their survival models. In this paper, I present an approach to modeling spatially autocorrelated random effects in survival data.

Specifically, I apply a Bayesian approach in which random effects at "neighboring" locations are allowed to exhibit spatial dependence (Banerjee, Wall, and Carlin 2003; Banerjee and Carlin 2003; Banerjee, Carlin, and Gelfand 2004). (The definition of neighbors is generalizable and need not imply contiguity). This spatial dependence is incorporated by specifying a conditionally autoregressive (CAR) prior developed by Besag, York, and Mollie (1991) for application in Bayesian image analysis. I employ the CAR prior to allow for spatially autocorrelated random effects in time-to-event data across neighboring units, with the neighbors defined via an adjacency matrix.

Table 1 lists the models I examine in this paper. I examine both semiparametric Cox and parametric (Weibull) survival models, and examine both unit-specific (individual) and hierarchical (shared) frailty models. In all, I examine the performance of eight different models: standard Cox and Weibull models with no frailties, Cox and Weibull models with shared non-spatial frailties, Cox and Weibull models with shared spatial frailties, and Weibull models with unit-specific non-spatial frailties and with unit-specific spatial frailties. I apply these models to Box-Steffensmeier, Arnold, and Zorn's (1997) data on the timing of U.S. House members' position announcements on the North American Free Trade Agreement.

The spatial shared frailty models outperform standard non-frailty models and non-spatial frailty models in both the semiparametric and parametric analyses. The hierarchical spatial

Weibull model also outperforms the unit-specific spatial Weibull model. Posterior means for substantive covariates differ between the spatial and non-spatial models. These results argue for the importance of modeling spatial dependence in random effects survival models.

The paper is structured as follows. The next section discusses survival models and the standard approach to frailty modeling in which the random effects are treated as independent. Next, I examine the Bayesian spatial survival modeling approach in which spatial dependence between neighboring random effects is modeled with a spatial prior. I next apply Bayesian spatial survival modeling to the timing of U.S. House members' position announcements on the North American Free Trade Agreement. The following section compares and contrasts the results of the Bayesian survival models and discusses their implications for our understanding of spatial effects in legislative position taking. I conclude by discussing potential future extensions of spatial frailty modeling within political science.

2 Survival Models

Survival models seek to explain how the risk, or hazard, of an event occurring at a given time is affected by covariates of theoretical interest. In a single event analysis, such as this paper's, the hazard rate is the instantaneous risk of a unit experiencing the event at a given time given that it has survived (i.e., not experienced the event) up to that time.[2] A critical distinction in survival analysis is how the baseline hazard (the hazard of the event in the absence of any covariate effects, i.e., the time dependency in the event process) is parameterized. In the semi-parametric Cox model, no parametric distribution is specified for the baseline hazard. As a consequence, rather than employing a specific distribution for the intervals between event occurrences, the Cox model incorporates information only for the observed event times. In contrast, in parametric survival models such as the Weibull or

[2]Single event analysis is contrasted with multiple events analysis, in which units are at risk of experiencing multiple events of substantive interest.

Gompertz, a specific parametric form is assumed for the underlying baseline hazard.[3]

In choosing between the Cox model and its parametric alternatives, one faces a tradeoff between the Cox model's flexibility (to various shapes of the baseline hazard) vs. the parametric models' more precise estimates of duration dependency (if the correct parametric distribution is chosen) and capacity for out-of-sample prediction. The semi-parametric Cox model and its principal parametric alternatives, however, share the common assumption of proportional hazards: covariates are assumed to have proportional effects on the baseline hazard that do not change with time. If this assumption is true, then as Box-Steffensmeier and Zorn (2001, 973) note, "the effects of covariates are constant over time."

The proportional hazards assumption is consequential because assuming that hazards are proportional when, in fact, they are non-proportional, can produce biased parameter estimates and decreases in the power of significance tests (Box-Steffensmeier and Zorn 2001, 972). It is thus critical to test for the proportionality of hazards (via, for example, Grambsch and Therneau's global test for nonproportional hazards and Harrell's rho test for covariate-specific non-proportionality) and model non-proportionality when it occurs (typically by interacting covariates that violate the proportional hazards assumption with a logged measure of time) (see Box-Steffensmeier and Jones 2004).

I examine both semi-parametric and parametric spatial survival models. The parametric analysis examines the Weibull model, which is frequently employed by researchers interested in parametric survival analysis.

In the Cox model, the hazard rate is:

$$h(t_i; \mathbf{x}_i) = h_0(t_i)\exp(\beta'\mathbf{x}_i) \tag{1}$$

[3]The Cox model is referred to as semi-parametric because although no distributional form is assumed for the baseline hazard, the risk of event occurrence is still parameterized as a function of covariates.

where t_i is the time to event or censoring for unit i, h_0 is the baseline hazard, $\mathbf{x}_i$ is a vector of covariates, and β is a vector of parameters. The Cox model, unlike parametric survival models, includes no intercept because the baseline hazard is not parameterized (Box-Steffensmeier and Jones 2004, 49). In the Weibull model, the hazard rate is:

$$h(t_i; \mathbf{x}_i) = \rho t_i^{\rho-1} \exp(\beta' \mathbf{x}_i) \tag{2}$$

where ρ is a shape parameter for the baseline hazard, β now includes an intercept term (as the baseline hazard is modeled using the Weibull distribution), and the remaining notation is as in (1). The shape parameter, ρ, reflects the shape of the monotonic hazard in the Weibull model, with $\rho > 1$ reflecting a monotonically rising hazard rate, $\rho < 1$ reflecting a monotonically declining hazard, and $\rho = 1$ reflecting a flat hazard (Box-Steffensmeier and Jones 2000, 25).

3 Standard Frailty Models

The Cox and Weibull models in (1) and (2) assume that factors affecting the hazard of event occurrence are included in the covariate vector, $\mathbf{x}_i$. What is the effect of omitting factors that affect the hazard? As Box-Steffensmeier and Jones (2004, 147) note, omitting such factors reduces the effect of covariates in the model that increase the hazard and increases the effect of covariates that reduce the hazard. Thus, scholars will wish to have a way to account for covariates excluded from the model that affect the hazard rate.

One strategy for accounting for such omitted covariates is the inclusion of random effect, or frailty terms. The frailty terms account for the fact that some units are at greater risk of experiencing the event of interest, that is, are more frail, due to factors not incorporated in the model. Either of two standard frailty approaches are adopted, depending upon the researcher's prior beliefs about the nature of the unobserved heterogeneity in risk propensity. If the researcher believes that units exhibit their own unique frailties, she will incorporate unit-specific, or individual frailty terms for each unit in her data. Alternatively, if she believes that units are clustered such that units within the same cluster share the same frailty while frailties are independent across clusters, she will incorporate hierarchical, or shared frailty terms for each cluster in her data. The basic structure of the two standard frailty modeling approaches is similar, and I thus motivate my discussion with the individual frailty model, discussing afterward how this approach is modified for the case of shared frailties.

3.1 Individual and Shared Frailty Models with Independent Random Effects

The hazard rate in the Cox model with standard independent individual frailties takes the form:

$$h(t_i; \mathbf{x}_i) = h_0(t_i)\exp(\beta'\mathbf{x}_i + W_i), \tag{3}$$

while the hazard rate in the Weibull model with standard independent individual frailties takes the form:

$$h(t_i; \mathbf{x}_i) = \rho t_i^{\rho-1}\exp(\beta'\mathbf{x}_i + W_i), \tag{4}$$

where $W_i \equiv \log \omega_i$ is the individual frailty term, the remaining notation in (3) is as in (1), and in (4) as in (2). As can be seen from the equations, the hazard in the frailty model is a function not only of the covariate vector, $\mathbf{x}_i$, but also of the random effect, W_i.

If unmodeled factors produce significant heterogeneity in risk propensity, the variance of the frailties will be distinguishable from zero. Thus, estimating whether random effects should be included in the model involves specifying a probability distribution for the frailties and estimating the variance, θ, of the frailties. The gamma and inverse Gaussian are often chosen for the random effects distribution (Therneau and Grambsch 2000, 232-234).

The individual frailty modeling approach is appropriate only if the researcher believes that each unit has a unique unmodeled frailty. If, however, the researcher believes that units are clustered in a hierarchical structure, such that units within the same cluster share a common frailty, a hierarchical, shared frailty modeling approach is appropriate instead.

Here the Cox and Weibull shared frailty models take the form:

$$h(t_{ij}; \mathbf{x}_{ij}) = h_0(t_{ij})\exp(\beta'\mathbf{x}_{ij} + W_j), \tag{5}$$

and

$$h(t_{ij}; \mathbf{x}_{ij}) = \rho t_{ij}^{\rho-1}\exp(\beta'\mathbf{x}_{ij} + W_j), \tag{6}$$

where unit i is now nested in cluster or stratum j, and the individual frailty, W_i, is now replaced by a shared frailty, $W_j \equiv \log \omega_j$, for units nested in stratum j. Inference regarding the appropriateness of the shared frailty model proceeds analogously to the individual frailty case. A probability distribution is specified for the shared frailty terms, and a variance, θ, that is distinguishable from zero indicates that there are unmodeled shared risk factors.

Both the individual and shared frailty approaches make a critical assumption. In both approaches, the random effects are assumed to be independent. In the individual frailty model, each unit has a unique frailty that is independent of other individual random effects. In the shared frailty model, units within the same cluster share a common frailty, but the frailties are assumed independent across these higher-level units.

I argue that this assumption of independent random effects will often be unrealistic in political science data. Theories of political event processes predict spatial dependence in event occurrence. If we are unable to model fully this spatial dependence via substantive covariates, this will produce spatially autocorrelated random effects. In the individual frailty approach, neighboring units will have spatially dependent frailties. In the shared frailty approach, neighboring strata will have spatially dependent frailties. Given evidence of spatial autocorrelation in many political science data (see, e.g., O'Loughlin, Flint, and Anselin 1994; Shin and Ward 1999; Gimpel and Cho 2004; Beck, Gleditsch, and Beardsley 2006; Darmofal 2006a), I argue that scholars will often wish to model spatial dependence in random effects.

The modeling approach I examine here also provides a more realistic and flexible approach to modeling dependent data than standard hierarchical models. The standard approach makes a knife-edge assumption: units in the same strata are assumed to exhibit dependence, but the strata themselves, even neighboring strata, are assumed to be independent. Thus, for example, in a shared frailty model of legislative behavior with state-level strata (a commonly employed choice for clustering political science data), Democratic Representatives Jerrold Nadler of New York's 8th Congressional District and Robert Menendez of New Jersey's neighboring 13th Congressional District would be treated as spatially independent in the 103rd Congress. This is not consistent with our understanding of how spatial proximity promotes legislative interaction (see Caldeira and Patterson 1987). More generally, because social science processes are inherently social and interdependent, spatial dependence often does not stop at the stratum's edge. Instead, neighboring strata often exhibit spatial dependence. As a consequence, scholars will often wish to incorporate spatial dependence between strata rather than making the knife-edge assumption of spatial independence across strata.

In short, political scientists often have reason to expect spatial dependence in the political event processes they seek to model. Standard approaches to modeling these event processes to date, however, have treated observations as spatially independent. The Bayesian spatial survival modeling approach I examine provides researchers with an approach for modeling the spatial autocorrelation that is predicted by our theories of event processes in political science. In the next section, I examine semiparametric and parametric Bayesian approaches for modeling spatially dependent survival data.

4 Bayesian Spatial Survival Modeling

Typically, spatial data in political science take the form of lattice data, in which a continuous spatial surface is divided into a grid of (typically irregular) lattice objects, or polygons, such as counties, states, congressional districts, or the like.[4] The critical step that distin-

[4]Geostatistical data are less commonly employed in political science. In contrast to polygonal lattice data, geostatistical data are sample data from a continuous spatial surface. For an application of Bayesian spatial survival modeling to geostatistical data, see Banerjee,

guishes spatial modeling of event processes from standard modeling approaches for event processes is the incorporation of adjacency information for the observations and the parameterization of spatial dependence across neighboring polygons. From a Bayesian perspective, this involves incorporating a prior to account for the spatial dependence in the hazards. Typically, a conditionally autoregressive (CAR) prior incorporating adjacency information is employed to model this spatial dependence.

Before examining the CAR prior, it is critical to distinguish what we mean by "adjacency" and "neighboring locations." Often, substantive theory suggests that spatial dependence operates geographically; in this case, the term "neighboring" can be taken literally, with spatial dependence modeled for adjacent polygons. As Beck, Gleditsch, and Beardsley (2006) demonstrate, however, dependence in other applications may take a non-spatial form (they model dependence as a function of trade flows between countries). Thus, "neighbors" need not imply contiguity and indeed, there need not be any inherent spatial component to the analysis at all. As a consequence, the modeling approach examined here is quite general, and can be applied to the more general question of correlated hazards across observations.

Neighbors are defined via a weights matrix, $\mathbf{A}$. In an unnormalized weights matrix such as that employed in this paper's NAFTA application, each neighbor of a unit is given a weight of 1, while each non-neighbor of a unit is given a weight of 0. (Thus, $a_{ii'} = 1$ if units i and i' are neighbors, and $a_{ii'} = 0$ if units i and i' are non-neighbors.) This spatial weights approach differs from the standard i.i.d. perspective on random effects, which, in a Bayesian framework, implies an exchangeable prior with a (non-spatial) weights matrix. In the exchangeable prior, all non-diagonal elements of the non-spatial weights matrix are given a common value, such as 1. In such an approach, the random effects are exchangeable under any geographic permutation of the data (Bernardinelli and Montomoli 1992, 988).

In the spatial CAR modeling approach, the definition of neighbors via the weights matrix

Wall, and Carlin 2003. Because lattice data are much more common than geostatistical data in political science, I do not consider geostatistical data further in this paper.

is a critical step in modeling spatially dependent event processes, since it delimits the possible spatial dependence that may be identified. The definition of neighbors should thus be guided by substantive theory. Because prior theory (e.g., Caldeira and Patterson 1987) argues that spatial proximity between legislators is likely to affect the timing of position announcements, I examine spatial dependence across geographically adjacent locations in my application.

Once the researcher has defined neighbors via the weights matrix, this information is then incorporated in the CAR prior. In recent years, biostatisticians have employed the CAR prior for Bayesian spatial shared frailty modeling (see Banerjee, Wall, and Carlin 2003; Banerjee and Carlin 2003; Banerjee, Carlin, and Gelfand 2004). I extend these models for the first time to the case of Bayesian spatial individual frailty models. I first examine the CAR priors for the individual and shared spatial cases and then examine the full semiparametric and parametric specifications incorporating the CAR prior.

4.1 The Conditionally Autoregressive (CAR) Prior

The standard, non-spatial frailty models in equations 3-6 all assume that the random effects or frailty terms are independent. From a Bayesian perspective, this is consistent with a specification in which the random effects distribution is conditional on a hyperparameter, λ, with an exchangeable prior, where λ refers to the precision (the inverse variance, i.e., the inverse of θ) of the random effects distribution. As in the case of the CAR prior, the λ prior is a unidimensional precision prior for the joint distribution of the random effects vector. The single dimensional prior for the random effects distribution is employed in survival modeling research by, e.g., Banerjee, Wall, and Carlin (2003), Banerjee, Carlin, and Gelfand (2004), and Lawson (forthcoming). As stated earlier, the exchangeable prior is induced by not distinguishing between neighboring and non-neighboring units in the weights matrix, but instead treating both neighbors and non-neighbors as exchangeable.[5]

[5]This section draws on the presentation and notation in Bernardinelli and Montomoli 1992, Banerjee, Wall, and Carlin 2003, Banerjee and Carlin 2003, and Banerjee, Carlin, and

The exchangeable prior, however, is likely to be problematic for many survival modeling applications. As the well-known Galton's problem recognizes, neighboring units are likely to share similar risk propensities, due either to behavioral diffusion or to shared risk factors. If we are unable to model fully the sources of risk propensity, neighboring units will share spatially autocorrelated frailties. As a consequence, we will often wish to relax the assumption of exchangeability. This is accomplished by allowing the precision parameter, λ, to reflect a conditionally autoregressive prior that incorporates neighbor definitions via the spatial weights matrix.

In the spatial individual frailty model, this CAR(λ) prior has a joint distribution proportional to:

$$\lambda^{(I-G)/2}\exp\left[-\frac{\lambda}{2}\sum_{i \text{ adj } i'}(W_i - W_{i'})^2\right] \propto \lambda^{(I-G)/2}\exp\left[-\frac{\lambda}{2}\sum_{i=1}^{I} m_i W_i(W_i - \overline{W}_i)\right], \qquad (7)$$

where I is the number of units in the data, G is the number of unconnected (island) units, i adj i' indicates that units i and i' are adjacent, $\overline{W}_i$ is the average of the $W_{i'\neq i}$ neighboring W_i, and m_i is the number of adjacencies (Bernardinelli and Montomoli 1992; Banerjee, Wall, and Carlin 2003, 126).

The conditional distribution of the spatial random effects that results from the CAR prior is then:

$$W_i|W_{i'\neq i} \sim N(\overline{W}_i, 1/(\lambda m_i)). \qquad (8)$$

By incorporating the spatial locations of units, the CAR prior thus produces a conditional distribution for the random effects that is normally distributed with a conditional mean equal to the average of the random effects for neighbors of i, and a conditional variance that is inversely proportional to the number of units neighboring i (Thomas et al. 2004). Thus, where the exchangeable prior displaces the random effect estimates toward a global mean by not distinguishing between neighbors and non-neighbors, the spatial CAR prior displaces

---

Gelfand 2004.

these estimates toward a local mean (Bernardinelli and Montomoli 1992, 989).

The CAR prior for the spatial shared frailty model follows accordingly. In the spatial shared frailty model, the individual unit i is now nested in a higher-level cluster or stratum j, and the random effect refers to this higher-level stratum, W_j. The CAR(λ) prior for the spatial shared frailty model has a joint distribution proportional to:

$$\lambda^{(J-H)/2}\exp\left[-\frac{\lambda}{2}\sum_{j \text{ adj } j'}(W_j - W_{j'})^2\right] \propto \lambda^{(J-H)/2}\exp\left[-\frac{\lambda}{2}\sum_{j=1}^{J} m_j W_j(W_j - \overline{W}_j)\right], \quad (9)$$

where J is the number of higher-level strata in the data, H is the number of unconnected (island) strata, j adj j' indicates that strata j and j' are adjacent, $\overline{W}_j$ is the average of the $W_{j'\neq j}$ neighboring W_j, and m_j is the number of higher-level adjacencies (Bernardinelli and Montomoli 1992; Banerjee, Wall, and Carlin 2003, 126).

The resulting conditional distribution for the strata-level spatial random effects is then:

$$W_j|W_{j'\neq j} \sim N(\overline{W}_j, 1/(\lambda m_j)). \quad (10)$$

Analogous to the individual frailty case, the CAR prior thus produces a conditional distribution for the spatial shared frailties that is normally distributed with a conditional mean equal to the average of the random effects for strata neighboring stratum j, and a conditional variance that is inversely proportional to the number of strata neighboring j (Thomas et al. 2004). The individual and shared spatial frailty models also require that a hyperprior, $p(\lambda)$, be assigned to λ. Generally, a $Gamma(a,b)$ hyperprior is chosen (Banerjee, Carlin, and Gelfand 2004). A reference prior should also be employed to gauge the effect of the Gamma hyperprior (see, e.g., Gelman 2006, Gelman et al. 2004).

The CAR prior is an improper prior, with the mean of the distribution of the spatial random effects undefined. Any constant can be added to the random effects and the prior remains unchanged (Banerjee, Carlin, and Gelfand 2004, 80). As a consequence of its impropriety, the CAR model can only be used as a prior and not as a likelihood (Banerjee, Carlin, and Gelfand 2004, 80). Because the CAR prior, as a pairwise-difference prior, is identified only up to an additive constant, a constraint must be imposed on the frailties to identify an intercept term (Besag et al. 1995; Banerjee, Wall, and Carlin 2003, 126). To identify an intercept in the Weibull models, I thus impose the constraint that the frailties sum to zero.

4.2 Semiparametric Cox Models with Spatial Frailties

For the Bayesian semiparametric Cox model, the joint posterior distribution is:

$$p(\beta, \mathbf{W}, \lambda | \mathbf{t}, \mathbf{x}, \gamma) \propto L(\beta, \mathbf{W}; \mathbf{t}, \mathbf{x}, \gamma) p(\mathbf{W}|\lambda) p(\beta) p(\lambda), \tag{11}$$

where $\mathbf{t}$ is the collection of event times, γ is the collection of event indicators, and the remaining notation is as in previous equations. The first term on the right in (11) is the Cox likelihood and the remaining terms are the CAR distribution of the frailties, the priors on β, and the hyperprior on λ. The likelihood for the Bayesian Cox model with spatial individual frailties is then[6]:

$$Ł(\beta, \mathbf{W}; \mathbf{t}, \mathbf{x}, \gamma) \propto \prod_{i=1}^{I} \{h_0(t_i; \mathbf{x}_i)\}^{\gamma_i} \exp\{-H_0(t_i) \exp(\beta' \mathbf{x}_i + W_i)\}, \tag{12}$$

while the likelihood for the Bayesian Cox model with spatial shared frailties is:

$$Ł(\beta, \mathbf{W}; \mathbf{t}, \mathbf{x}, \gamma) \propto \prod_{j=1}^{J} \prod_{i=1}^{n_j} \{h_0(t_{ij}; \mathbf{x}_{ij})\}^{\gamma_{ij}} \exp\{-H_0(t_{ij}) \exp(\beta' \mathbf{x}_{ij} + W_j)\}. \tag{13}$$

In contrast to standard Cox frailty models, the inclusion of the conditionally autoregressive prior in the Cox spatial frailty models incorporates the potential spatial dependence among frailties at neighboring locations. The individual and shared frailty Cox models are completed

[6]The Cox model with spatial individual frailties, like its non-spatial counterpart, is only identified in the presence of time-varying covariates. My NAFTA application does not include time-varying covariates and thus I do not estimate Cox models with individual frailties.

by assigning appropriate priors for β and λ.

4.3 Parametric Weibull Models with Spatial Frailties

The joint posterior distribution for the Bayesian parametric Weibull model is:

$$p(\beta, \mathbf{W}, \rho, \lambda | \mathbf{t}, \mathbf{x}, \gamma) \propto L(\beta, \mathbf{W}, \rho; \mathbf{t}, \mathbf{x}, \gamma) p(\mathbf{W}|\lambda) p(\beta) p(\rho) p(\lambda), \tag{14}$$

where the notation is as in (11), except that ρ, the shape parameter for the baseline hazard in the Weibull, is now included. The first term on the right is now the Weibull likelihood, the second is again the CAR distribution of the random effects, and the remaining terms are the remaining prior distributions.

The likelihood for the Weibull model with spatial individual frailties is proportional to:

$$\prod_{i=1}^{I} \{\rho t_i^{\rho-1} \exp(\beta' \mathbf{x}_i + W_i)\}^{\gamma_i} \exp\{-t_i^{\rho} \exp(\beta' \mathbf{x}_i + W_i)\}, \tag{15}$$

while the likelihood for the Weibull model with spatial shared frailties is proportional to:

$$\prod_{j=1}^{J} \prod_{i=1}^{n_j} \{\rho t_{ij}^{\rho-1} \exp(\beta' \mathbf{x}_{ij} + W_j)\}^{\gamma_{ij}} \exp\{-t_{ij}^{\rho} \exp(\beta' \mathbf{x}_{ij} + W_j)\}. \tag{16}$$

As in the Cox spatial frailty specification, the Weibull spatial frailty model differs from the standard Weibull model in its inclusion of the CAR prior, incorporating spatial dependence among neighboring frailties. The individual and shared spatial Weibull specifications are then completed by assigning appropriate priors for β, ρ, and λ. Generally, a $Gamma(\alpha, 1/\alpha)$ prior is chosen for ρ and, as in the Cox model, a $Gamma(a, b)$ prior is chosen for λ.

4.4 Models with Both Spatial and Non-Spatial Frailties

Researchers may also be interested in estimating survival models with both spatial and non-spatial frailties. Such an approach can be useful in examining the relative contributions of spatial and non-spatial effects. Care must be taken here, however, as the spatial and independent random effects are now identified only through their priors (Banerjee and Carlin 2003, 526). The likelihood for the Cox model with shared spatial and non-spatial frailties is:

$$Ł(\beta, \mathbf{W}; \mathbf{t}, \mathbf{x}, \gamma) \propto \prod_{j=1}^{J} \prod_{i=1}^{n_j} \{h_0(t_{ij}; \mathbf{x}_{ij})\}^{\gamma_{ij}} \exp\{-H_0(t_{ij}) \exp(\beta' \mathbf{x}_{ij} + W_j + V_j)\}, \tag{17}$$

while the likelihood for the Weibull model with shared spatial and non-spatial frailties is:

$$Ł(\beta, \mathbf{W}, \rho; \mathbf{t}, \mathbf{x}, \gamma) \propto \prod_{j=1}^{J} \prod_{i=1}^{n_j} \{\rho t_{ij}^{\rho-1} \exp(\beta' \mathbf{x}_{ij} + W_j)\}^{\gamma_{ij}} \exp\{-t_{ij}^{\rho} \exp(\beta' \mathbf{x}_{ij} + W_j + V_j)\}, \tag{18}$$

where V_j represents the non-spatial frailties (with $V_j \stackrel{iid}{\sim} N(0, 1/\tau)$) and the rest of the notation remains as it is in (13) and (16) respectively (Banerjee and Carlin 2003, 526). The individual model with spatial and non-spatial frailties is defined analogously. The Cox model with spatial and non-spatial frailties is then completed with the choice of priors for β, λ, and τ, and the Weibull is completed with priors for β, λ, ρ, and τ. The priors previously discussed for λ and ρ are generally employed for the joint spatial/non-spatial case, while τ is generally given a $Gamma(c,d)$ prior (Banerjee and Carlin 2003).[7]

5 Application of Spatial Frailty Modeling to the Timing of Position Taking on NAFTA

I apply Bayesian spatial frailty modeling to the timing of position announcements by members of the U.S. House of Representatives on the North American Free Trade Agreement (NAFTA). In their analysis of position timing on NAFTA, Box-Steffensmeier, Arnold, and Zorn (1997) incorporated spatial influences via a dummy variable indicating whether the member's district shared a land border with Mexico. Spatial effects in position announce-

[7] I attempted to estimate a Weibull model with spatial and non-spatial shared frailties for the NAFTA data, but was unable to achieve convergence for several parameters.

ments, however, were unlikely to be limited to the border's edge. I posit, instead, that members from neighboring locations were likely to announce positions at similar times, and thus exhibit spatial dependence in position timing. Consistent with Galton's problem, two sets of factors were likely to produce this spatial autocorrelation.

On the one hand, spatial dependence in position timing may have occurred as a result of behavioral diffusion. Caldeira and Patterson (1987) demonstrate that members from neighboring legislative districts are more likely to develop friendships with each other than are members from more spatially distant districts. Accordingly, I expect more frequent – and more effective – interpersonal interaction among members from neighboring districts than among members from more distant districts, and greater similarity in the timing of legislative position announcements as a result.

Alternatively, spatially proximate members may announce positions at similar times despite little or no communication between each other. Members from neighboring locations are more likely to share similar constituencies, and similar constituent concerns, than more spatially distant members. Thus, similar factors that lead to cue-taking and cue-giving among same-state Senators (Matthews and Stimson 1975) may also produce spatial dependence in the timing of position announcements among neighboring House members. Members from neighboring locations may also share similar partisan, ideological, or demographic characteristics, producing similarity in both policy positions and in the timing of announcements of these positions. Thus, even if members from neighboring districts rarely talk, they may still exhibit spatially dependent position timing due to common district or personal attributes.

My analysis thus contrasts with many previous analyses of legislative behavior that have focused on dependence *within* same-state delegations in that I allow for dependence across state boundaries. In addition to potential spatial effects, I model three sets of factors likely to affect position timing. These three sets of factors are constituency, institutional, and individual (member) influences (see Box-Steffensmeier, Arnold, and Zorn 1997).[8]

---

[8]I chose the covariate profile for illustrative purposes, to examine the effects of unmodeled

I model constituency factors with two covariates (full descriptions of all covariates in the models, from Box-Steffensmeier, Arnold, and Zorn (1997, 336), are provided in the Appendix, along with descriptive statistics). NAFTA carried different policy effects for constituents with different economic profiles. Specifically, NAFTA was expected to pose significant dislocation effects on union members and low-income citizens, with the high-paying jobs of the former and the low-skill jobs of the latter particularly threatened by foreign competition. As a result, I include the covariates, *Union Membership*, measuring the percentage of private-sector workers in the member's district who were union members, and *Household Income*, measuring the district's median household income. As Box-Steffensmeier, Arnold, and Zorn (1997, 327) note, these covariates as coded in their analysis and mine do not present clear expectations for effects on the timing of NAFTA announcements. In each case, members with high or low values on the variable are expected to announce earlier than members with intermediate levels due to clearer signals of policy preferences from constituents. I retain Box-Steffensmeier, Arnold, and Zorn's operationalization of these covariates.

Institutional influences in the House of Representatives are measured with three covariates. *NAFTA Committee* is a dichotomous measure indicating whether the member was on a committee that acted on NAFTA implementing legislation. Because committee membership provides an effective stage for cue-giving to other members, NAFTA committee members are expected to announce positions earlier than non-committee members. *Republican Leadership* and *Democratic Leadership* are dummy variables indicating whether the member held a position in his or her party's leadership in the House. Because Republican leaders were united in their support for NAFTA, a Republican leadership position is expected to be associated with earlier position taking. In contrast, Democratic leaders were divided on the trade agreement, cross-pressured in some cases by opposition to the agreement among constituents and support for the agreement by the Clinton administration. As a consequence, this covariate

---

and modeled spatial dependence on frailty and substantive covariate estimates and model choice statistics.

does not present clear expectations for the timing of members' announcements.

Finally, two interaction terms incorporating member ideology are included in the model to capture individual-level influences. The two covariates are *Ideology*Union Membership* and *Ideology*Household Income*.[9] Box-Steffensmeier, Arnold, and Zorn (1997) measure ideology with members' Chamber of Commerce voting scores (purged of the NAFTA vote) on economic issues. Members with voting scores > 50 (indicating more pro-business voting records) are scored 1 on the dichotomous measure, while members with scores at or below 50 are scored 0. Box-Steffensmeier, Arnold, and Zorn posit that it is the interaction of member ideology and district ideology (as proxied by union membership and household income) that is most relevant for the timing of NAFTA position announcements. I follow their specification and incorporate these two interaction terms in the model. Due to potential cross-pressures between member and district ideology, neither interaction term presents clear expectations regarding effects on position timing.

5.1 Neighbor Definitions and Priors

I employ distinct neighbor definitions for the individual and hierarchical frailty models. For the individual frailty model, I created an adjacency matrix with a queen contiguity definition, in which each district contiguous to member i's district in the 103rd Congress is a neighbor of member i and each district that is not contiguous to member i's district is a non-neighbor. For the hierarchical model, I nest members of Congress within states and allow for spatial dependence across the state-level random effects. Again I employ a queen contiguity neighbor definition, in which each state contiguous to state i is a neighbor of state i and each state not contiguous to state i is a non-neighbor of state i.[10]

[9]As Box-Steffensmeier, Arnold, and Zorn (1997, 17) note, *Ideology* is not entered separately in the model because doing so would produce an intercept in the Cox model.

[10]The data thus differ in the individual and shared frailty analyses. Rep. Don Young (R-AK) is excluded from both analyses, since his district is not contiguous to any other districts.

Employing these two distinct conceptions of neighbors has particular utility for examining the validity of the standard approach to modeling dependence in legislative behavior in which neighbors are nested within states. The standard approach corresponds to the non-spatial shared frailty model with independent state-level random effects. The shared spatial frailty model allows us to examine the validity of treating the state-level random effects as independent. The individual spatial frailty model, in turn, allows us to examine whether spatial dependence should be modeled more locally via neighboring districts, including those in neighboring states, or less locally via spatially autocorrelated state-level frailties.

I complete the specifications by specifying appropriate priors for the parameters in the models. Given that spatial frailty models have not previously been employed in political science and prior substantive research provides little information regarding the values of the spatial random effects or the values of substantive covariates in the presence of these spatial frailties, I prefer vague prior distributions, relying on the data to overwhelm the priors. I employ a vague hyperprior for λ of $Gamma(.01,\ .01)$, a prior of $N(0, .001)$ for β_0 in the Weibull model, and priors of $N(0, .00001)$ for the remaining β in both the Cox and Weibull models.[11] I set $\alpha = .01$, producing a $Gamma(.01,\ 100)$ prior for the shape parameter, ρ, in

Rep. Neil Abercrombie (D-HI) and Rep. Patsy Mink (D-HI) are included in the individual frailty analysis because their districts are contiguous to each other but are excluded from the hierarchical frailty analysis because Hawaii is not contiguous to any other states.

[11]To examine the sensitivity of the results to the gamma CAR prior, I also estimated the models using a uniform reference prior for the CAR prior, following Gelman (2006) and Gelman et al. (2004). The results are similar whether the gamma or uniform priors are used. Although the posterior means for the frailty variance parameter, θ, are somewhat smaller under the uniform prior (0.191 vs. 0.195 in the spatial Cox shared frailty models, .005 vs. .011 in the spatial Weibull individual frailty models, and .007 vs. .018 in the spatial Weibull shared frailty models), in all cases, the 95 percent Bayesian credible intervals for this parameter are distinguishable from zero. The estimates for the other parameters exhibit

the Weibull model.[12]

I employ Markov Chain Monte Carlo techniques to characterize the posterior densities of the parameters and hyperparameters of interest.[13] Specifically, I employed Gibbs sampling for two separate Markov chains with overdispersed starting values of 0 and 1 for the intercept, β_0, in the Weibull models, $\pm$ 3 standard errors from the frequentist Cox estimates in Box-Steffensmeier, Arnold, and Zorn (1997, 331) for the remaining β, .01 and .1 for ρ, .001 and 1 for λ, and .01 and .1 for c and r. I employed 5,000 burn-in iterations for each Markov chain. Convergence was diagnosed via Gelman and Rubin's diagnostic (Gill 2002, 399-402), with the diagnostic indicating convergence for each parameter in each model. I retained 10,000 post burn-in iterations for each chain, providing a sample size of 20,000.

6 Spatial Dependence in the Timing of Position Taking on NAFTA

I first examine how the risk of a U.S. House member announcing a position on NAFTA varied as a function of time.[14] The empirical baseline hazard for the Cox model is non-

only marginal changes. The reference prior estimates are available from the author.

[12]I employ an Andersen-Gill counting process formulation for the Cox model. The counting process requires the specification of two additional priors for c, the researcher's degree of confidence in her belief regarding the underlying hazard function, and r, the researcher's prior regarding the failure rate per unit of time. I express weak priors regarding both the values of the hazard function and the failure rate via priors on c and r of (.0001, .00001) and (.001, .0001), respectively. For additional information on the counting process approach, see Andersen-Gill 1982, Clayton 1991, and Spiegelhalter et al. 2003.

[13]WinBUGS 1.4.1 was used for the Bayesian analysis.

[14]I retain Box-Steffensmeier, Arnold, and Zorn's (1997, 330) coding and assume that members came under risk of announcing a position on August 12, 1992, the day that Rep. Peter Visclosky (D-IN) announced his opposition to NAFTA. "Undecided" and "leaning" positions are not included in this measure. Members who did not make a public announcement of their

monotonic, but generally increasing over time. Member announcements on NAFTA were backloaded – more than 90 percent of announcements occurred more than 300 days after Rep. Peter Visclosky's (D-IN) announcement of his opposition to NAFTA on August 12, 1992. The data are, moreover, heavily clustered. More than 80 percent of members announced their positions on September 9, 1993 or later; nearly half of members announced their positions in the month leading up to the House vote.

Such heavily clustered data mitigate against finding spatial dependence. With so many members announcing their positions concurrently, it is clear that non-spatial effects played a significant role in the timing of position taking. Thus, my particular application serves as a conservative test of spatial dependence in time-to-event data. If we find evidence of significant spatial effects even in survival data marked by large spikes in event occurrence such as those in the timing of NAFTA position announcements, we will have reason to expect even stronger spatial effects in less temporally clustered data.

6.1 Assessing Model Choice

Because the spatial CAR prior is an improper prior, Bayes factors cannot be used to choose among the alternative models (e.g., Gill 2002). Instead, I use the Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002) to assess model choice across the spatial and non-spatial Cox and Weibull survival models. The DIC, like the more familiar Akaike Information Criterion (AIC), combines measures both of model fit and of the effective number of parameters (the latter component penalizes models that overfit the data).

The deviance statistic is central to the model fit component of the DIC. The deviance

---

position prior to the House vote on H.R. 3450, the North American Free Trade Agreement (NAFTA) Implementation Act, on November 17, 1993 are recorded as announcing their position on this date. Full descriptions of the data can be found in Box-Steffensmeier, Arnold, and Zorn (1997).

statistic takes the form:

$$D(\theta) = -2 \log f(\mathbf{y}|\theta) + 2 \log h(\mathbf{y}), \tag{19}$$

where, as Banerjee and Carlin (2003, 532) note, $f(\mathbf{y}|\theta)$ is the likelihood for the observed data given the parameter vector θ and $h(\mathbf{y})$ is a function of only the data.[15] The intuition behind the deviance statistic is to examine the improvement in fit produced by the estimation of the parameter vector θ. The model fit is then summarized using the posterior expectation of the deviance, $\overline{D} = \mathrm{E}_{\theta|y}[D]$.

Because the estimation of unnecessary parameters in the parameter vector θ naturally improves model fit, it is important to penalize for overfitting the model. This is done by calculating the effective number of parameters, p_D, for the model. The effective number of parameters reflects the relative role that the data play in estimating the parameters vs. the priors, with larger estimates of the effective number of parameters indicating that the data play a larger role. As Gelman et al. (2004, 182) note, in calculating the effective number of parameters, a parameter receives a value of 1 if it is estimated from the data alone with no input from the prior, a value of 0 if it is estimated from the prior alone with no input from the data, and an intermediate value between 1 and 0 depending upon the relative contributions of the data and the prior. The effective number of parameters is calculated as:

$$p_D = E_{\theta|y}[D] - D(E_{\theta|y}[\theta]) = \overline{D} - D(\overline{\theta}) \tag{20}$$

where $\overline{D}$ is, again, the posterior expectation of the deviance and $D(\overline{\theta})$ is the deviance taken at the posterior expectations (Banerjee, Wall, and Carlin 2003, 127). The effective number of parameters is thus the deviance of the posterior means subtracted from the posterior mean of the deviance (Spiegelhalter, et al. 2003). Combining the measure of model fit with the

[15]This section draws on the discussion and notation in Banerjee, Wall, and Carlin (2003).

penalty for overfitting, the DIC then takes the form:

$$DIC = \overline{D} + p_D \tag{21}$$

The DIC of models fit to the same data can be compared to determine the appropriate model choice. As with other information criteria, smaller values of the DIC are favored over larger values.[16]

Table 2 reports the effective number of parameters and DIC values for three Cox models and five Weibull models: standard Cox and Weibull models with no random effects, Cox and Weibulls with non-spatial shared (state-level) random effects, Cox and Weibulls with spatial shared (state-level) random effects, a Weibull model with non-spatial individual (district-level) random effects, and a Weibull model with spatial, individual (district-level) random effects. In each model, the specification included the covariates discussed in Section 5.

As Table 2 shows, in each of the three comparisons between the spatial and non-spatial frailty models, the spatial frailty models outperform their non-spatial counterparts, as indicated by the smaller DIC values. There is, in short, spatial dependence in the timing of NAFTA announcements that is not fully captured by the substantive covariates in the model. Treating the random effects as though they were spatially independent, as we typically do, reflects model misspecification.

Importantly, the information criterion advantages are not produced by overfitting the models with additional parameters. In both the case of the Cox model and the shared frailty Weibull model, the spatial frailty model has a smaller effective number of parameters (p_D) than does the non-spatial frailty model. These spatial frailty models thus enjoy a parsimony

[16]As Banerjee and Carlin 2003, 532) note, this can be seen from the fact that small values of the posterior expectation of the deviance reflect a good fit while a small number of effective parameters reflects parsimony. The goal, as with any information criterion, is thus to combine model fit and parsimony.

advantage over their non-spatial counterparts. In the third comparison, for the individual frailty Weibull models, the spatial model has only a very marginal increase in the effective number of parameters over its non-spatial counterpart (12.33 vs. 12.17).

Examining the DICs for the various models, clear patterns emerge. The models that incorporate spatial dependence in state-level frailties are the preferred model in both the semiparametric Cox and parametric Weibull cases. The Cox model with spatial shared frailties outperforms both the standard Cox model and the Cox model with non-spatial shared frailties. The Weibull with spatial shared frailties outperforms the standard Weibull, the Weibull with non-spatial shared frailties, and the two Weibull individual frailty specifications. Whether considering a semiparametric or parametric modeling approach, in this case scholars should fit a model that accounts for spatial dependence across state-level effects rather than fitting either a standard model that doesn't account for unmodeled heterogeneity in risk propensity, or a frailty model that treats this heterogeneity as spatially independent.

More broadly, scholars modeling legislative behavior have become accustomed to clustering legislators by state and treating members from different states as independent, conditional on the covariates. The DIC values in Table 2 question the validity of such an approach. Scholars, instead, should consider the possibility that members from neighboring states share common unmeasured characteristics that impact the behavior of interest. The results also argue that modeling heterogeneity via unit-specific random effects is not ideal either. The DICs indicate that the argument that the "uniqueness" of each individual actor precludes conceptual generalization is not valid for this particular case of legislative behavior. House members are not independent actors; neighboring legislators share common risk factors, whether due to direct behavioral interaction or shared attributes.

6.2 Cox and Weibull Results

I examine spatial dependence in the timing of NAFTA position announcements as well as its effects on substantive covariates via summaries of the posterior densities from the Bayesian

Cox and Weibull analyses.[17] Table 3 presents the summaries for the semiparametric Cox MCMC analysis, while Table 4 presents the summaries for the parametric Weibull MCMC analysis. In both tables, the first cell entry is the mean of the posterior density of the particular parameter of interest while the cell entry in parentheses below is the corresponding 95% Bayesian credible interval (formed by taking the 2.5 and 97.5 posterior percentiles). Descriptions of the models in each column are provided below the tables.

Examining the Cox summaries in Table 3, it is clear that a standard, non-spatial frailty model understates the unmodeled heterogeneity in the data. The posterior mean of the variance of the random effects, θ, is more than twice as large in the spatial model (column 3) as in the non-spatial model (column 2). Spatially proximate members share common unmodeled risk factors that distinguish them from their spatially distant colleagues. Modeling these risk factors as though they were spatially independent markedly understates the unmodeled heterogeneity in risk propensity.

By mapping the frailties from the non-spatial and spatial Cox models, we can further see the problems that are induced by modeling spatially dependent risk factors as though they were spatially independent. Figure 1 presents a map of the posterior means of the non-spatial state-level frailties – means estimated under the assumption of spatial independence. Figure 1 suggests a checkerboard pattern, with little spatial clustering in the random effects.

Figure 2 maps the posterior means from the spatial Cox model that takes into account the spatial dependence between the state-level frailties. As can be seen from Figure 2, there is, in fact, a strong spatial clustering in the unobserved risk factors. There are distinct spatial bands in the random effects. Portions of the Northeast, upper Plains, and West were marked by particularly high risk propensity (and thus, all else equal, members from these states were more likely to be early announcers of NAFTA positions). Members from the Rocky Mountain states shared the next level of hazards. Next, members from a band of states extending from

[17] The Grambsch and Therneau global test and Harrell's rho covariate-specific tests showed no violations of the proportional hazards assumption.

the industrial Midwest, Southwest into Oklahoma shared similar risk factors. Next, we see a set of shared hazards extending from the border states to Pennsylvania. Finally, we see a clustering in the Carolinas of low risk factors for NAFTA announcements. In contrast to the false impression of independent random effects suggested by Figure 1, we can see from Figure 2 that the spatial location of members played a significant role in the timing of their position announcements on NAFTA.

Table 3 demonstrates the importance of modeling spatial dependence in random effects if we wish to draw accurate inferences about other substantive covariates of interest. The posterior means for the spatial frailty model differ from those in the standard, non-frailty Cox model, and in all but one case, differ more from the latter than do the means from the non-spatial frailty model. Note, for example, the changes in the posterior means for *Union Membership* and the *Ideology*Union Membership* interaction in the spatial model vs. the standard Cox model. Similarly, the means for the *Democratic Leadership* effects are noticeably different across the two models. Where the standard Cox model predicts that a position in the Democratic leadership increases the hazard of a NAFTA announcement by 9.5 percent, the spatial Cox model predicts an increase in the hazard of only 1.8 percent.

Table 4 reports the posterior summaries for the five Weibull models. Although the parametric assumption of the Weibull makes it more restrictive, the Weibull specifications also allow for the estimation of individual frailty models. As a result, the Weibulls allow us to compare how spatial frailty effects differ at the individual and shared levels.

In both the individual and shared Weibulls, the frailty effects are larger in the spatial models than in the non-spatial models. The mean of the variance parameter, θ, however, is noticeably larger in the Weibull with spatial shared frailties than in the Weibull with spatial individual frailties. Thus, consistent with the DICs, it is particularly important in estimating position timing to account for spatial dependence across state-level random effects. As we would expect given the smaller values of θ in the Weibull models, the differences in effects across models on substantive covariates are not as dramatic as in the Cox specifications.

Examining the effects of substantive factors on NAFTA position timing, we can gauge the effect of constituency, institutional, and individual factors by examining the posterior summaries from the best-performing model according to the DIC, the Cox spatial shared frailty model. As we can see from Table 3, the main effect of large union memberships in a member's district was to increase the hazard of a NAFTA position announcement. Economically conservative members from districts with large union memberships (alternatively, economically liberal members from districts with low union memberships), however, had reduced hazards of NAFTA announcements, as indicated by the negative value for the mean on the *Ideology*Union Membership* interaction. This suggests, then, that cross-pressures delayed announcements for some members. Also of note, members of the Republican leadership had increased hazards of NAFTA announcements; this is as we would expect given the Republican leadership's united support for the trade agreement. Overall, constituency, institutional, and individual characteristics all influenced the timing of NAFTA position announcements.

7 Conclusion

In this paper I have presented an approach to modeling spatial dependence in political event processes. The importance of modeling spatial autocorrelation in survival data is clear: many of our theories of event processes in political science predict spatial dependence among neighboring units. If we are unable to model fully this spatial dependence, the result will be spatially autocorrelated unmeasured risk factors among neighboring units. Political scientists examining event processes will, therefore, often not wish to assume that frailties among neighboring observations are spatially independent. Instead, scholars will often have strong theoretical justification for modeling spatial dependence in the random effects among neighboring observations. The conditionally autoregressive prior allows political scientists to incorporate this spatial dependence in their survival models.

The paper's results highlight the importance of modeling the spatial autocorrelation that is common to so many political science data. The Deviance Information Criterion (DIC) values favored the spatial shared frailty models (in both semiparametric and parametric forms) over both standard non-frailty models and non-spatial frailty models. The posterior mean of the random effects variance parameter in both models, moreover, was noticeably larger than the mean for the non-spatial variance parameter. Accounting for this unmodeled spatial dependence produced distinct changes in the posterior means for substantive covariates.

This initial analysis of spatial frailties in political time-to-event data examined single-spell event processes. Future research can extend Bayesian spatial survival models to examine spatial dependence in more complex event processes. Often political events are repeated events; units are at risk of experiencing the same type of event multiple times. Substantive theory predicts significant spatial dependence in repeated event processes. For example, international conflicts occur among the same participants repeatedly, with particular regions such as the Middle East particularly prone to such repeated conflicts. A more complex model would examine spatial dependence in competing risks, in which units at any time are at risk of experiencing any of multiple types of events. We could easily, for example, imagine policymakers making choices among various health policy reforms where the choice of the particular policy is shaped by the decision processes in neighboring states or communities.

At their base, political concerns are shared concerns. As a corollary, political events carry significance, in large part, because they are not isolated events. A conflict in one location may spill over and produce conflicts in neighboring locations. Democratization in one nation may produce a wave of democratization in neighboring countries. Political event processes, in short, take place in both space and time. To draw valid inferences about the factors shaping political event processes we must account for both these spatial and temporal dimensions. The Bayesian approach examined here presents an effective approach for political scientists wishing to account for space and time in their models of political event processes.

Table 1: Model Comparisons

Model	Cox	Weibull
Standard	X	X
Non-Spatial Individual Frailties		X
Spatial Individual Frailties		X
Non-Spatial Shared Frailties	X	X
Spatial Shared Frailties	X	X

Table 2: Model Choice Statistics

Model	P_D	DIC
Standard Cox	77.43	5201.97
Cox with Non-Spatial Shared Frailties	98.92	5170.03
Cox with Spatial Shared Frailties	95.45	5166.85
Standard Weibull	9.24	5204.68
Weibull with Non-Spatial Individual Frailties	12.17	5228.10
Weibull with Spatial Individual Frailties	12.33	5219.19
Weibull with Non-Spatial Shared Frailties	12.41	5195.71
Weibull with Spatial Shared Frailties	11.77	5192.05

Table 3: Posterior Summaries for Cox Models

	1	2	3
Union Membership	3.512	3.424	2.736
	(1.875, 5.154)	(1.243, 5.543)	(0.393, 5.055)
Household Income	-0.041	-0.106	-0.132
	(-0.168, 0.087)	(-0.245, 0.027)	(-0.272, 0.009)
NAFTA Committee	-0.009	-0.023	-0.040
	(-0.160, 0.142)	(-0.177, 0.130)	(-0.193, 0.114)
Republican Leadership	0.348	0.407	0.418
	(-0.008, 0.678)	(0.040, 0.758)	(0.051, 0.764)
Democratic Leadership	0.091	0.033	0.018
	(-0.244, 0.405)	(-0.312, 0.354)	(-0.328, 0.339)
Ideology*Union Membership	-4.174	-3.873	-3.778
	(-6.676, -1.664)	(-6.491, -1.201)	(-6.429, -1.138)
Ideology*Household Income	0.142	0.128	0.145
	(-0.035, 0.317)	(-0.052, 0.307)	(-0.034, 0.324)
θ		0.082	0.195
		(0.026, 0.178)	(0.056, 0.442)

Cell entries are the posterior means, with 95% credible intervals in parentheses.

(1 = Standard Cox model, 2 = Cox model with non-spatial shared frailties, 3 = Cox model with spatial shared frailties)

Table 4: Posterior Summaries for Weibull Models

	1	2	3	4	5
Constant	-19.14	-18.59	-18.84	-18.83	-18.91
	(-21.24, -17.29)	(-20.09, -17.04)	(-20.35, -17.63)	(-20.43, -17.09)	(-20.61, -17.09)
Union Membership	1.154	1.102	0.966	1.021	1.006
	(-1.035, 3.290)	(-1.094, 3.212)	(-1.292, 3.238)	(-1.267, 3.266)	(-1.344, 3.271)
Household Income	-0.034	-0.034	-0.042	-0.041	-0.044
	(-0.203, 0.133)	(-0.207, 0.136)	(-0.210, 0.128)	(-0.213, 0.129)	(-0.216, 0.129)
NAFTA Committee	-0.011	-0.011	-0.007	-0.008	-0.008
	(-0.220, 0.196)	(-0.222, 0.194)	(-0.217, 0.199)	(-0.217, 0.200)	(-0.216, 0.197)
Republican Leadership	0.119	0.117	0.124	0.123	0.124
	(-0.392, 0.585)	(-0.397, 0.582)	(-0.386, 0.587)	(-0.391, 0.584)	(-0.386, 0.588)
Democratic Leadership	0.027	0.025	0.026	0.026	0.028
	(-0.438, 0.453)	(-0.456, 0.456)	(-0.439, 0.452)	(-0.443, 0.455)	(-0.446, 0.454)
Ideology*Union Membership	-1.593	-1.514	-1.350	-1.412	-1.434
	(-5.167, 1.962)	(-5.028, 2.003)	(-4.977, 2.180)	(-5.057, 2.182)	(-5.006, 2.187)
Ideology*Household Income	0.043	0.040	0.048	0.044	0.048
	(-0.194, 0.282)	(-0.197, 0.281)	(-0.192, 0.288)	(-0.200, 0.285)	(-0.193, 0.286)
ρ	3.173	3.081	3.123	3.121	3.134
	(2.869, 3.520)	(2.828, 3.330)	(2.925, 3.371)	(2.834, 3.383)	(2.835, 3.413)
θ		0.008	0.011	0.011	0.018
		(0.002, 0.021)	(0.002, 0.032)	(0.003, 0.030)	(0.003, 0.064)

Cell entries are the posterior means, with 95% credible intervals in parentheses.

(1 = Standard Weibull model, 2 = Weibull model with non-spatial individual frailties, 3 = Weibull model with spatial individual frailties, 4 = Weibull model with non-spatial shared frailties, 5 = Weibull model with spatial shared frailties)

Appendix: Variable Descriptions

Dependent Variable: *Timing of Position.* The number of days after August 11, 1992 until the member of Congress stated a yes or no position on NAFTA. Mean: 403.14, Standard Deviation: 70.16, Minimum: 1, Maximum: 463.

Independent Variables

Union Membership. Proportion of private-sector workers belonging to a union in the member's district, 1991-92. Data are from the *Current Population Survey.* Values on variable are mean-centered. Mean: .00, Standard Deviation: .06, Minimum: -.10, Maximum: .20.

Household Income. Median household income in the district in thousands of dollars. Data are from the *Almanac of American Politics, 103rd Congress.* Values on variable are mean-centered. Mean: -.01, Standard Deviation: .84, Minimum: -1.62, Maximum: 2.65.

NAFTA Committee. Dichotomous variable indicating whether member was on a committee that acted on NAFTA implementing legislation. Coded 1 if the representative was a member, 0 otherwise. Mean: .30, Standard Deviation: .46, Minimum: 0, Maximum: 1.

Republican Leadership. Dichotomous variable indicating whether member was in a Republican leadership position in the House. Coded 1 if minority leader, conference chair, vice-chair, secretary, minority whip, chief deputy whip, deputy whip, or assistant deputy whip, 0 otherwise. Mean: .04, Standard Deviation: .20, Minimum: 0, Maximum: 1.

Democratic Leadership. Dichotomous variable indicating whether member was in a Democratic leadership position in the House. Coded 1 if Speaker, majority leader, caucus chair, vice-chair, secretary, majority whip, floor whip, ex-officio whip, chief deputy whip, or assistant deputy whip, 0 otherwise. Mean: .05, Standard Deviation: .22, Minimum: 0, Maximum: 1.

Ideology (included in interaction terms). Dichotomous variable based on 1993 Chamber of Commerce voting score (with NAFTA vote purged). Coded 0 if rating was ≤ 50, 1 otherwise. Mean: .44, Standard Deviation: .50, Minimum: 0, Maximum: 1.

References

Andersen, Per Kragh, and R.D. Gill. 1982. "Cox's Regression Model for Counting Processes: A Large Sample Study." *The Annals of Statistics* 10: 1100-1120.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Kluwer: Dordrecht.

Banerjee, Sudipto, and Bradley P. Carlin. 2003. "Semiparametric Spatio-Temporal Frailty Modeling." *Environmetrics* 14: 523-535.

Banerjee, Sudipto, Melanie M. Wall, and Bradley P. Carlin. 2003. "Frailty Modeling for Spatially Correlated Survival Data, with Application to Infant Mortality in Minnesota." *Biostatistics* 4(1): 123-142.

Banerjee, Sudipto, Bradley P. Carlin, and Alan E. Gelfand. 2004. *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, FL: Chapman & Hall.

Beck, Nathaniel, Kristian Gleditsch, and Kyle Beardsley. 2006. "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy." *International Studies Quarterly* 50(1): 27-44.

Bernardinelli, L., and C. Montomoli. 1992. "Empirical Bayes Versus Fully Bayesian Analysis of Geographical Variation in Disease Risk." *Statistics in Medicine* 11: 983-1007.

Berry, Frances Stokes, and William D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84(2): 395-415.

Berry, Frances Stokes, and William D. Berry. 1992. "Tax Innovation in the States: Capitalizing on Political Opportunity." *American Journal of Political Science* 36(3): 715-742.

Berry, William D., and Brady Baybeck. 2005. "Using Geographic Information Systems to Study Interstate Competition." *American Political Science Review* 99(4): 505-519.

Besag, Julian, Jeremy York, and Annie Mollie. 1991. "Bayesian Image Restoration, with Two Applications in Spatial Statistics." *Annals of the Institute of Statistical Mathematics* 43: 1-59.

Besag, Julian, Peter Green, David Higdon, and Kerrie Mengersen. 1995. "Bayesian Computation and Stochastic Systems." *Statistical Science* 10(1): 3-41.

Box-Steffensmeier, Janet M., Laura W. Arnold, and Christopher J.W. Zorn. 1997. "The Strategic Timing of Position Taking in Congress: A Study of the North American Free Trade Agreement." *American Political Science Review* 91: 324-338.

Box-Steffensmeier, Janet M., and Suzanna De Boef. 2006. "Repeated Events Survival Models: The Conditional Frailty Model." *Statistics in Medicine* 25(20): 3518-3533.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press.

Caldeira, Gregory A., and Samuel C. Patterson. 1987. "Political Friendship in the Legislature." *Journal of Politics* 49(4): 953-975.

Carpenter, Daniel P. 2002. "Groups, the Media, Agency Waiting Costs, and FDA Drug Approval." *American Journal of Political Science* 46(3): 490-505.

Chiozza, Giacomo, and H.E. Goemans. 2004. "International Conflict and the Tenure of Leaders: Is War Still "Ex Post" Inefficient?" *American Journal of Political Science* 48(3): 604-619.

Cho, Wendy K. Tam. 2003. "Contagion Effects and Ethnic Contribution Networks." *American Journal of Political Science* 47(2): 368-387.

Clayton, David G. 1991. "A Monte Carlo Method for Bayesian Inference in Frailty Models." *Biometrics* 47(2): 467-485.

Colaresi, Michael. 2004. "When Doves Cry: International Rivalry, Unreciprocated Cooperation, and Leadership Turnover." *American Journal of Political Science* 48(3): 555-570.

Darmofal, David. 2006. "The Political Geography of Macro-Level Turnout in American Political Development." *Political Geography* 25(2): 123-150.

Franzese, Robert J., Jr., and Jude C. Hays. 2007. "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data." *Political Analysis* 15(2): 140-164.

Gelfand, A.E., and B.K. Mallick. 1995. "Bayesian Analysis of Proportional Hazards Models Built from Monotone Functions." *Biometrics* 51: 843-852.

Gelman, Andrew. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models." *Bayesian Analysis* 1(3): 515-533.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 2004. *Bayesian Data Analysis*, 2nd Ed. Boca Raton, FL: Chapman & Hall/CRC.

Gill, Jeff. 2002. *Bayesian Methods: A Social and Behavioral Sciences Approach*. Boca Raton, FL: Chapman & Hall/CRC.

Gimpel, James G., and Wendy K. Tam Cho. 2004. "The Persistence of White Ethnicity in New England Politics." *Political Geography* 23: 987-1008.

Hill, Daniel H., William G. Axinn, and Arland Thorton. 1993. "Competing Hazards with Shared Unmeasured Risk Factors." *Sociological Methodology* 23: 245-277.

Hougaard, Philip. 2000. *Analysis of Multivariate Survival Data*. New York: Springer-Verlag.

Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman, OK: University of Oklahoma Press.

Jones, Bradford S. 1994. *A Longitudinal Perspective on Congressional Elections*. Ph.D. Dissertation, State University of New York, Stony Brook.

Klein, John P., and Melvin L. Moeschberger. 1997. *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer-Verlag.

Lawson, Andrew B. Forthcoming. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology.* Boca Raton, FL: Chapman & Hall/CRC.

O'Loughlin, John, Colin Flint, and Luc Anselin. 1994. "The Geography of the Nazi Vote: Context, Confession, and Class in the Reichstag Election of 1930." *Annals of the Association of American Geographers* 84(3): 351-380.

Regan, Patrick M. 2002. "Third-Party Interventions and the Duration of Intrastate Conflicts." *The Journal of Conflict Resolution* 46(1): 55-73.

Shin, Michael, and Michael D. Ward. 1999. "Lost in Space: Political Geography and the Defense-Growth Trade-Off." *The Journal of Conflict Resolution* 43(6): 793-817.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit (with Discussion)." *Journal of the Royal Statistical Society, Series B* 64(4): 583-639.

Spiegelhalter, David J., Andrew Thomas, Nicola G. Best, and Dave Lunn. 2003. *WinBUGS User Manual Version, Version 1.4, January 2003*. Available at: http: www.mrc-bsu.cam.ac.uk/bugs/

Starr, Harvey. 1991. "Democratic Dominoes: Diffusion Approaches to the Spread of Democracy in the International System." *The Journal of Conflict Resolution* 35(2): 356-381.

Therneau, Terry M., and Patricia M. Grambsch. 2000. *Modeling Survival Data: Extending the Cox Model*. New York: Springer-Verlag.

Thomas, Andrew, Nicky Best, Dave Lunn, Richard Arnold, and David Spiegelhalter. 2004. *GeoBUGS User Manual, Version 1.2, September 2004*. Available at: http: www.mrc-bsu.cam.ac.uk/bugs/

Warwick, Paul. 1992. "Economic Trends and Government Survival in West European Parliamentary Democracies." *American Political Science Review* 86(4): 875-887.

Figure 1: Non-Spatial Cox State-Level Frailties

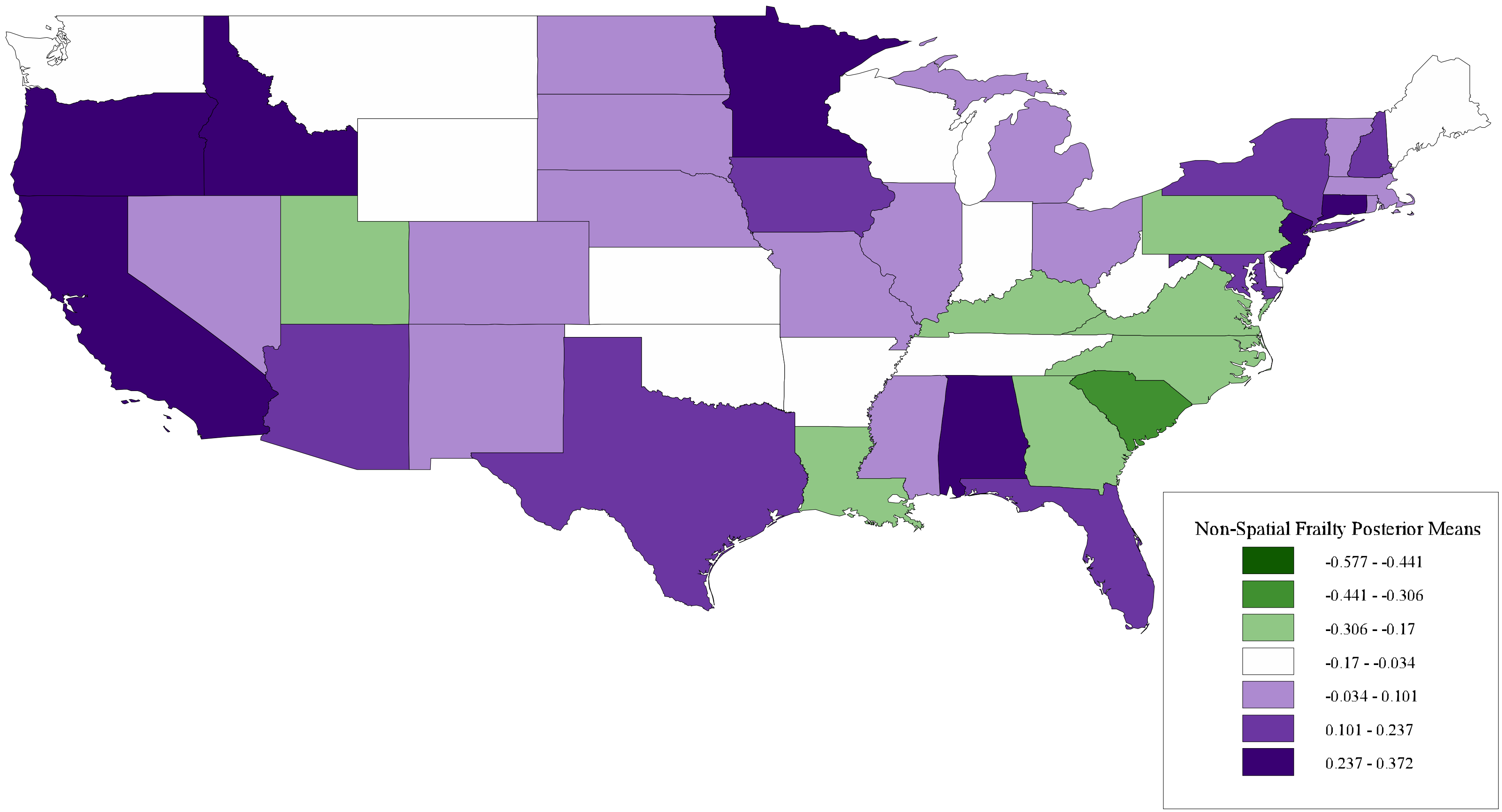

Figure 2: Spatial Cox State-Level Frailties